

---

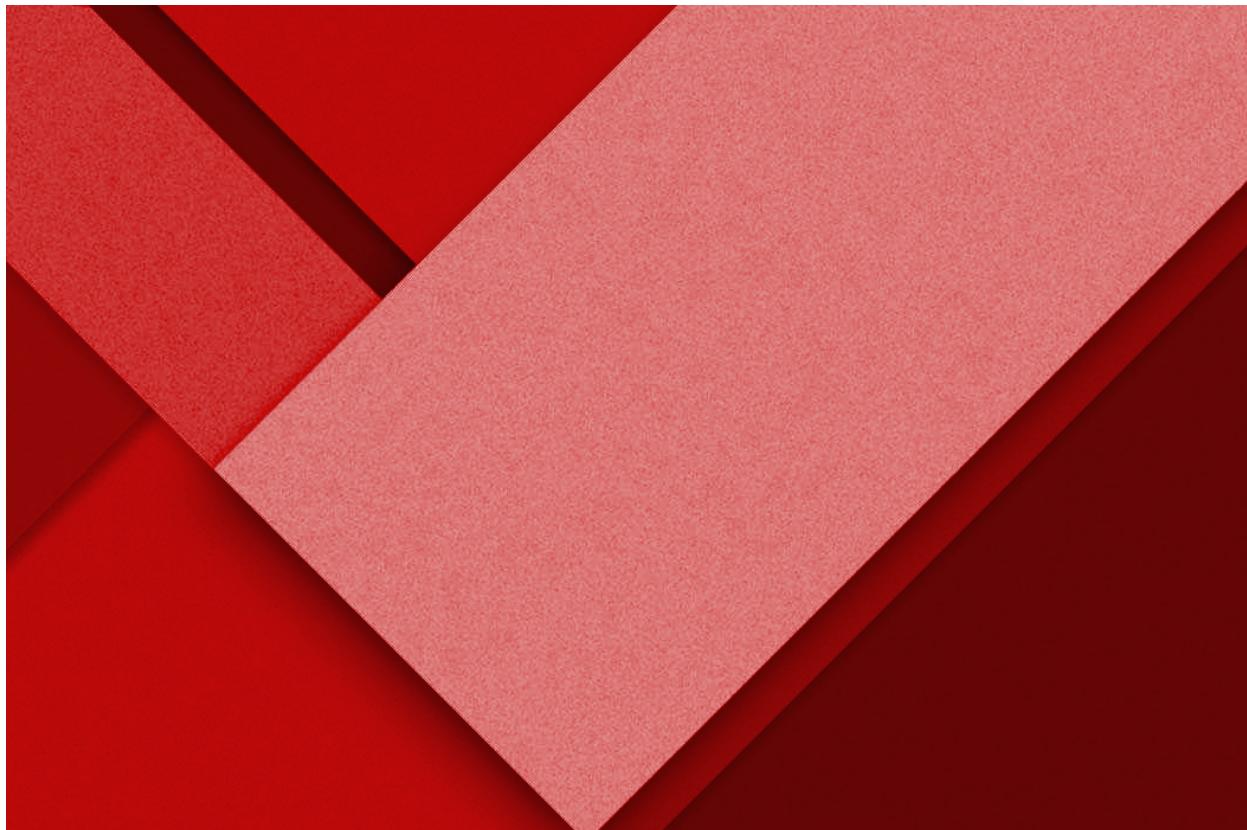
# CAHIER D'AVANCEMENT

YAKOUBI Oumayma

RIDOUAN Aafaf

LIROULET Inès

Master 2 IA



## **Avancement - Première Semaine : 03/12 → 10/12**

### **1. Choix de sujet**

Nous avons choisi de travailler sur l'analyse et la visualisation des comportements musicaux des utilisateurs de Spotify. Ce sujet met en avant l'importance de l'exploration des données personnelles pour fournir des insights significatifs et personnalisés.

### **2. Étude des solutions existantes**

Nous avons réalisé une recherche approfondie sur les outils et méthodes existants pour l'analyse des données Spotify. Cette étude a inclus l'exploration des fonctionnalités de D3.js pour les visualisations interactives, ainsi que des approches similaires pour comprendre les comportements musicaux.

### **3. Définition des questions du projet**

Nous avons identifié les principales questions auxquelles le projet doit répondre, notamment :

- Comment pourrait-on permettre aux utilisateurs de visualiser leurs morceaux préférés ?
- Comment pourrait-on permettre aux utilisateurs de visualiser leurs préférences musicales (genres de musique) ?
- Comment pourrait-on permettre aux utilisateurs d'analyser leur train de vie à partir de leurs horaires de consommation de musique ?
- Comment pourrait-on permettre d'analyser les humeurs d'un utilisateur selon le type de musique qu'il écoute ?
- Comment pourrait-on permettre aux utilisateurs de visualiser leur capacité à explorer de la nouvelle musique ?

### **4. Prototypage du projet:**

Nous avons commencé le prototypage en explorant les formats de données disponibles et en testant les premiers graphiques avec des bibliothèques comme D3.js. Cela nous a permis de définir une direction visuelle pour le projet.

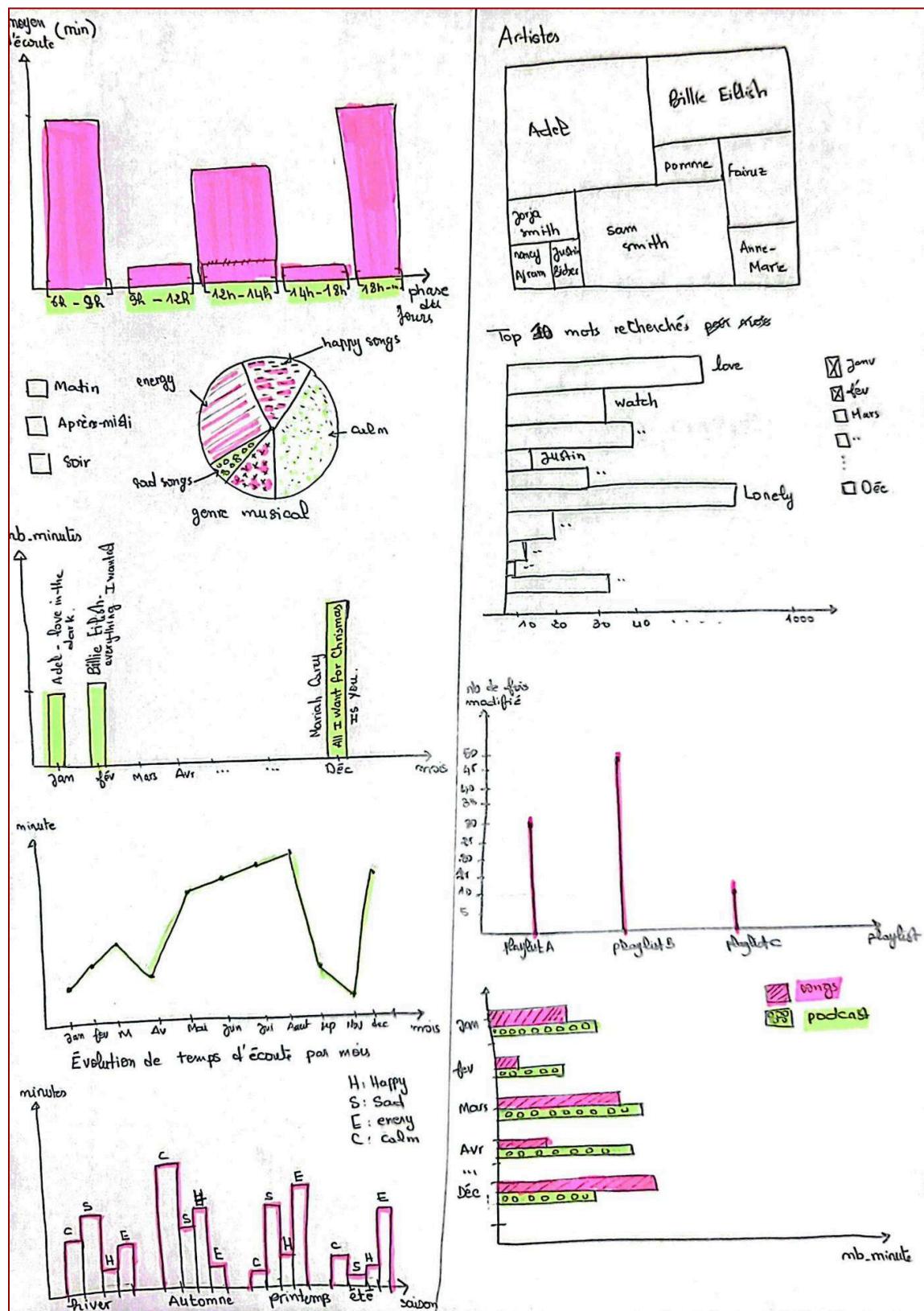
### **5. Production du document de cadrage**

Un document de cadrage a été produit, décrivant les objectifs, les délais, les besoins techniques, et les étapes clés pour mener à bien le projet. Ce document constitue une référence pour aligner les attentes des membres de l'équipe.

### **6. Création des premières esquisses du projet**

Nous avons conçu des esquisses préliminaires des visualisations principales. Ces esquisses incluent des concepts comme les graphes de relations artistes-morceaux, les

timelines d'écoute mensuelles, et les diagrammes illustrant les préférences musicales.



## Seconde semaine : 10/12 - 17/12

### 1. Compréhension et préparation des données

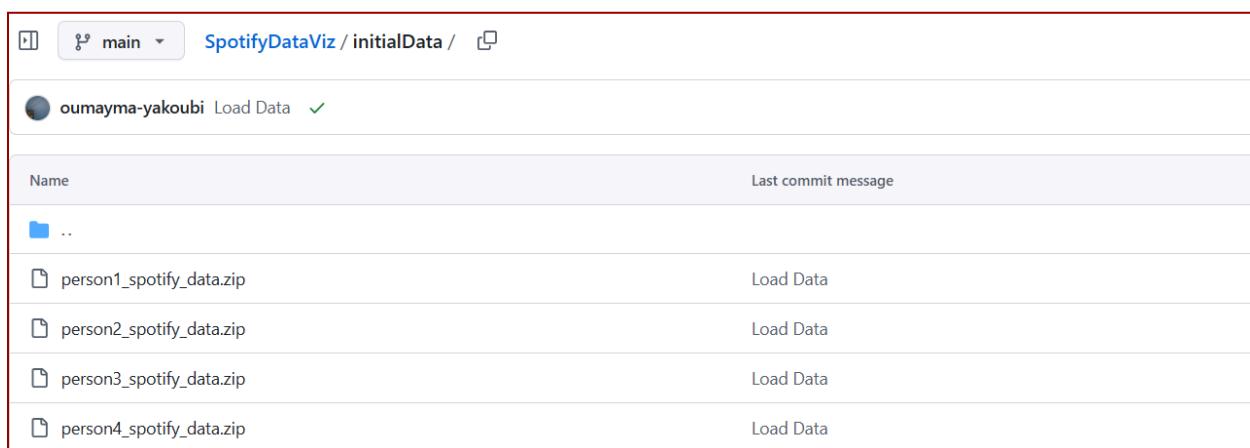
Le processus de compréhension et de préparation des données s'est déroulé en plusieurs étapes :

1.1 Exploration initiale : Nous avons examiné la structure des fichiers JSON fournis par Spotify, comprenant des informations telles que les chansons écoutées, les artistes favoris, les genres musicaux et les horaires d'écoute.

1.2- Nettoyage des données : Les fichiers ont été nettoyés pour éliminer les données redondantes ou inutiles pour notre étude.

1.3- Organisation : Nous avons structuré les données pour faciliter leur traitement. Nous avons mis l'ensemble des fichiers sur github et utiliser le lien vers ces fichiers pour accéder à ces données

1.4- Transformation des données : Les données brutes ont été transformées en un format adapté aux visualisations, en regroupant et en agrégeant les informations selon les besoins des différentes analyses prévues.



The screenshot shows a GitHub repository interface for 'SpotifyDataViz / index.json'. The code is as follows:

```

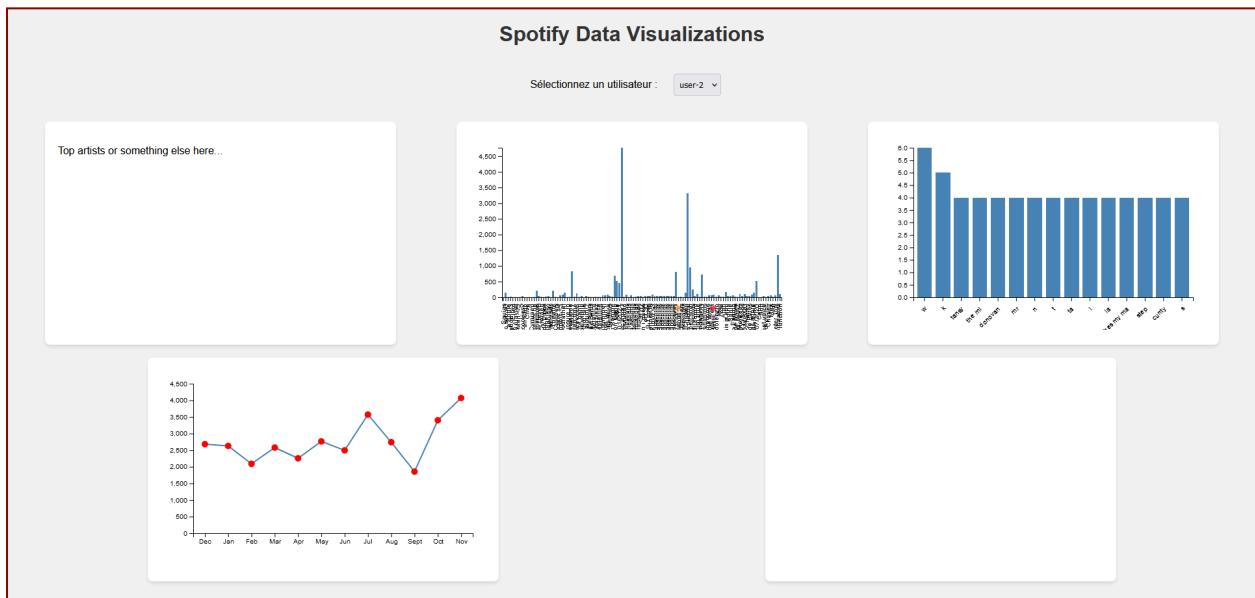
1   {
2     "user-1": [
3       "Identity.json",
4       "Playlist1.json",
5       "SearchQueries.json",
6       "StreamingHistory_music_0.json",
7       "StreamingHistory_podcast_0.json",
8       "Wrapped2023.json",
9       "YourLibrary.json"
10    ],
11    "user-2": [
12      "Identity.json",
13      "Playlist1.json",
14      "Playlist2.json",
15      "SearchQueries.json",
16      "StreamingHistory_music_0.json",
17      "StreamingHistory_music_1.json",
18      "StreamingHistory_podcast_0.json",
19      "Wrapped2023.json",
20      "Yourlibrary.json"
21    ],
22  ],

```

L'ensemble des données des utilisateurs est structuré dans un objet JSON permettant de les utiliser ensuite.

## 2. Mise en place d'une interface HTML rudimentaire

Une première interface a été mise en place pour permettre de voir les résultats des premières tentatives de visualisation. Cette interface est rudimentaire et temporaire.



### **3. Premières visualisations basiques**

Nous avons mis au point nos premières visualisations. Parmi elles :

- Une courbe représentant le nombre de minutes de musique écoutées par mois
- Un barplot représentant la taille des 20 plus grandes playlists
- Une treemap des 8 artistes les plus écoutés
- Un barplot des termes les plus recherchés

## **Troisième semaine : 17/12 - 24/12**

### **1. Autres visualisations basiques**

Deux autres ébauches de visualisations ont été mises au point :

- Un barplot de la quantité de musique écoutée selon les périodes de la journée
- Un barplot permettant de comparer l'écoute de musique par rapport à l'écoute de podcast selon les mois

### **2. Récupération des genres des chansons via la plateforme Spotify Developer**

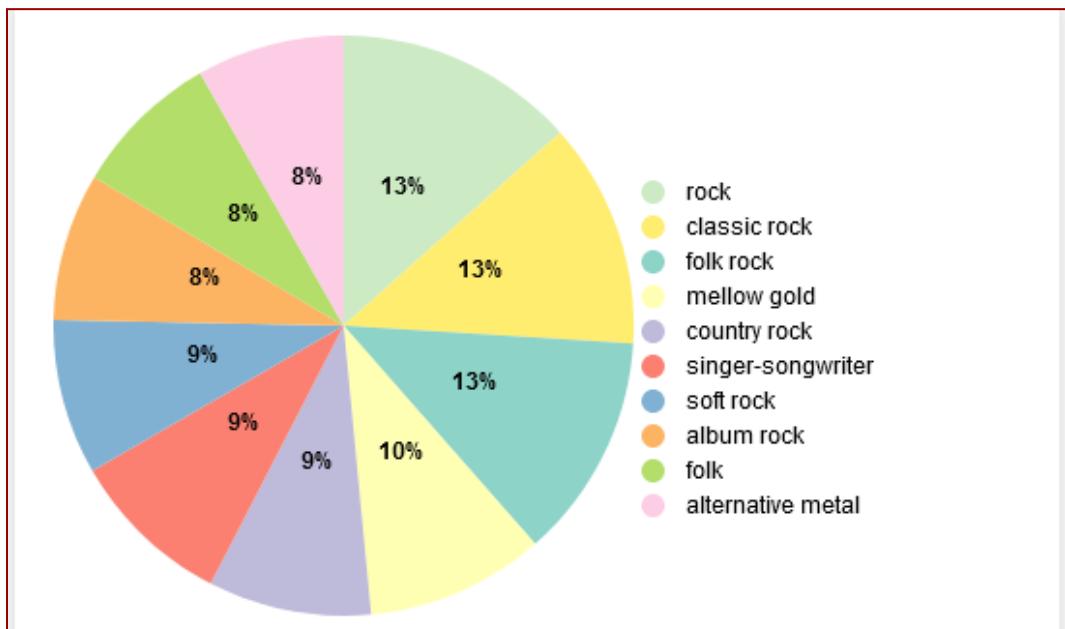
Pour l'extraction des genres musicaux, nous avons utilisé l'API Spotify Developer pour récupérer les informations sur les artistes présents dans les historiques d'écoute des utilisateurs. Voici les étapes principales du processus :

1. **Authentification** : Nous avons généré un jeton d'accès en utilisant les informations d'identification de l'application Spotify pour accéder à l'API.
2. **Extraction des artistes uniques** : À partir des historiques d'écoute, nous avons identifié les artistes uniques pour chaque utilisateur afin d'éviter les requêtes redondantes.
3. **Interrogation de l'API en lot** : Les artistes uniques ont été regroupés en petits lots, et chaque lot a été interrogé via l'API Spotify pour récupérer les informations, notamment les genres associés aux artistes.
4. **Gestion des limites de requêtes** : Une gestion des limites de taux (rate-limiting) a été mise en place pour respecter les contraintes de l'API Spotify, avec des mécanismes de temporisation et de reprise automatique en cas de dépassement.
5. **Sauvegarde des résultats** : Les données extraites, comprenant les genres musicaux associés à chaque artiste, ont été consolidées et sauvegardées dans des fichiers JSON pour chaque utilisateur.

MySpotifyDataViz / data / genre /	
 oumayma-yakoubi	data processing
Name	Last commit message
 ..	
 artistGenres_Aurelie.json	data processing
 artistGenres_Houssine.json	data processing
 artistGenres_Marouane.json	data processing
 artistGenres_Zakaria.json	update genre
 merged_genre_data.json	data processing

### 3. Ajout du pie chart de genres

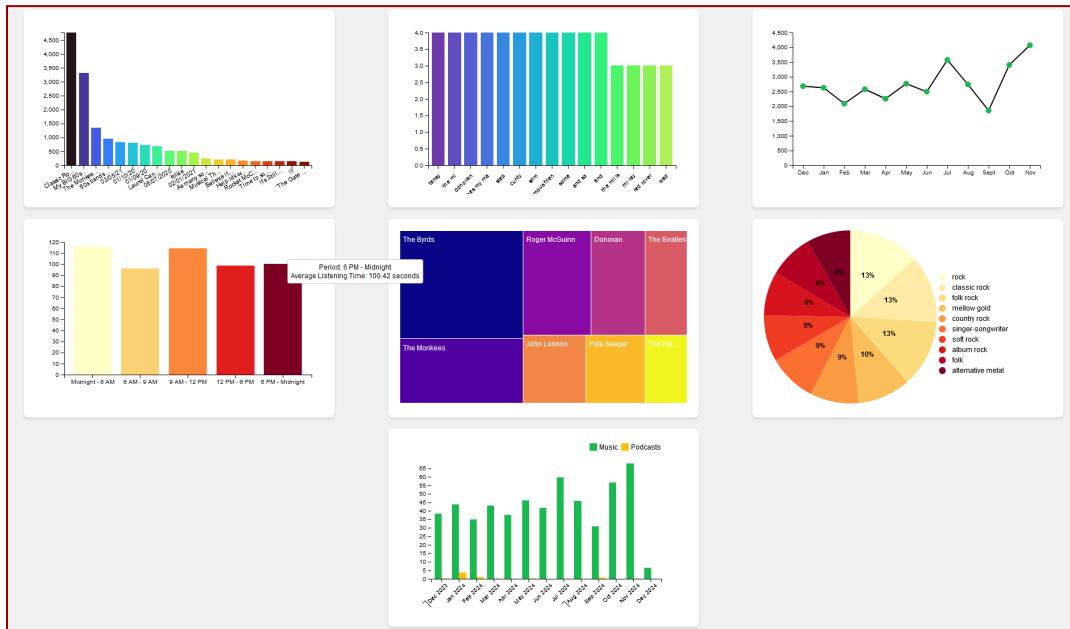
Après avoir récupéré l'information du genre auquel appartiennent les morceaux écoutés, une visualisation a pu être effectuée avec ces informations



## Quatrième semaine : 24/12 - 31/12

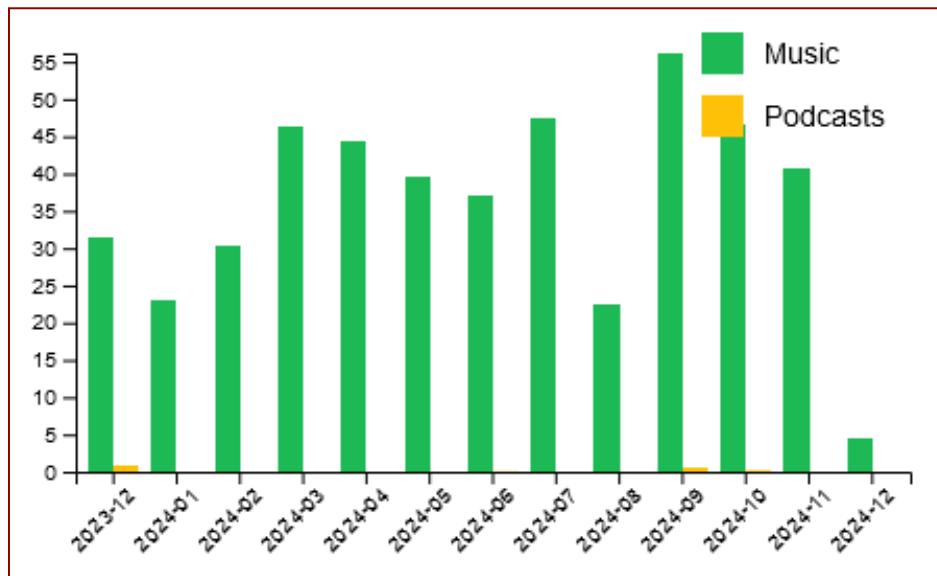
### 1. Essais de plusieurs palettes de couleur

Afin de commencer à saisir l'apparence finale du projet, plusieurs palettes de couleur ont été essayées pour nos visualisations, allant de palettes vertes pour rappeler le logo Spotify à des palettes complètement éloignées.

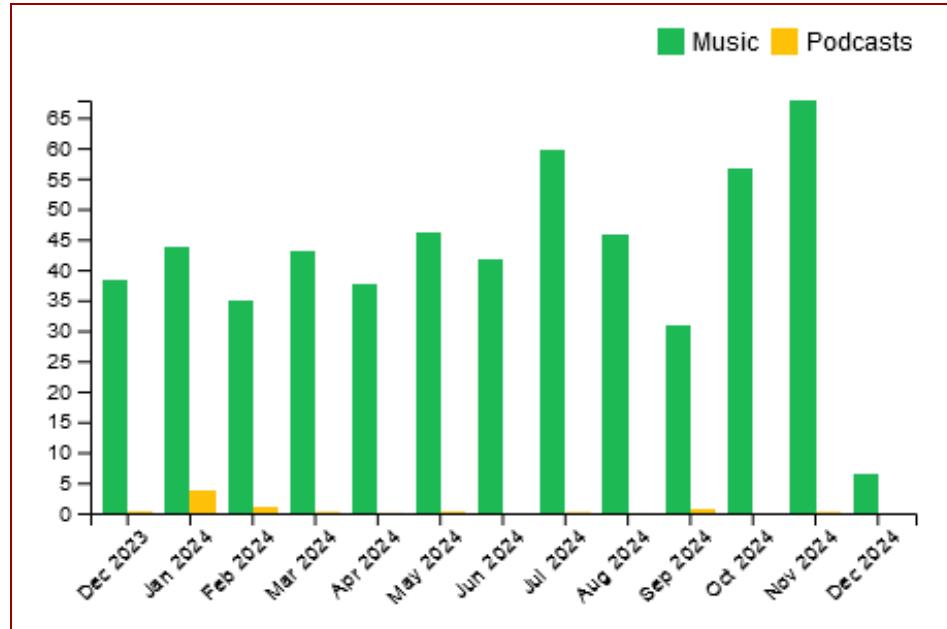


### 2. Ajustement de la légende music/podcast

La légende de la visualisation de différence entre écoute de musique et podcast a été fixée.



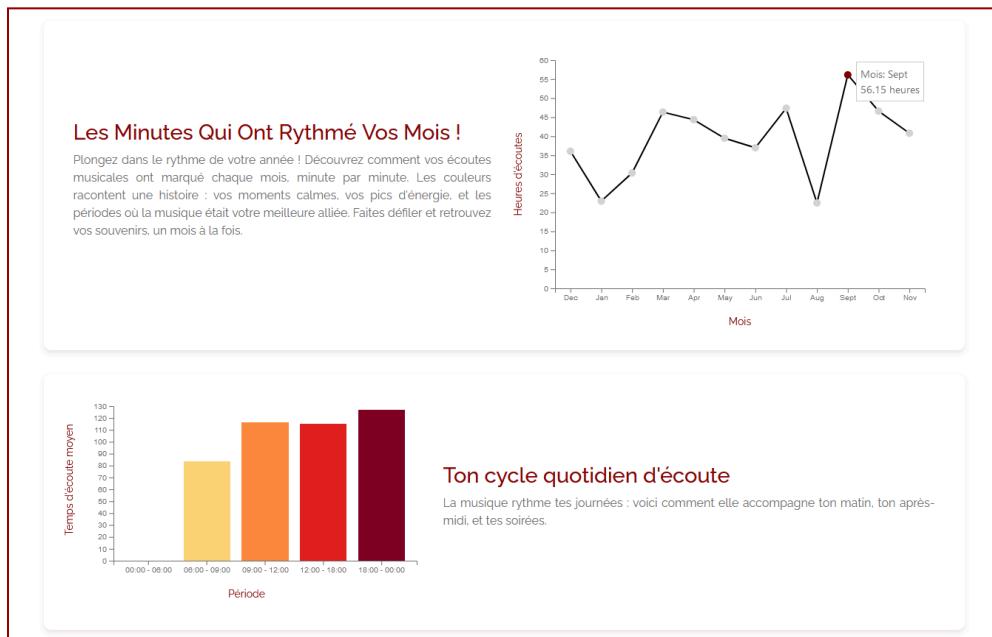
Légende avant



Après

### 3. Ajout de filtre entre les mois et périodes de la journée

Un filtre a été ajouté entre la visualisation de l'écoute par mois et de l'écoute selon les périodes de la journée. Ainsi, en cliquant sur le point représentant un mois dans la première visualisation, la seconde visualisation se met à jour pour montrer l'écoute moyenne selon les périodes uniquement sur le mois sélectionné.



## Cinquième semaine : 31/12 - 07/01

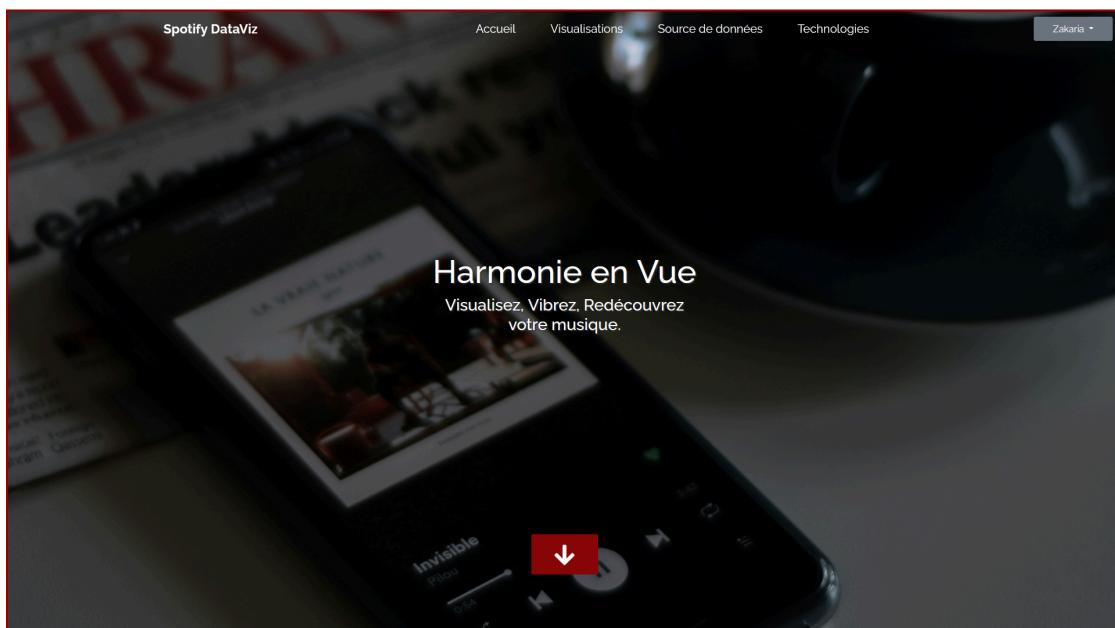
### 1. Décision de la palette de couleur

Nous avons finalement choisi une palette éloignée de celle de Spotify pour donner une dimension propre à notre projet.



### 2. Nouveau template

Un nouveau template a été mis au point pour le projet pour commencer la finalisation de l'apparence du projet.



### 3. Ajout de filtre mensuel sur les playlists

Un filtre a été ajouté pour permettre aux utilisateurs de visualiser l'évolution de leurs 20 plus grandes playlists de l'année selon les mois de l'année. La visualisation commence donc sur Décembre 2024, permettant de voir le nombre de morceaux présents dans la playlist à la fin de l'année, et un bouton permet de passer à un autre mois de l'année pour visualiser le nombre de morceaux que comportait cette playlist à ce moment-là.



## Sixième semaine : 07/01 - 14/01

### 1. Abandon de la visualisation top searches

La visualisation des termes les plus recherchés par l'utilisateur a finalement été abandonnée car trop peu pertinente. En effet, elle n'apportait pas grand chose à l'utilisateur et comportait des mots tronqués. En effet, l'interface de recherche de Spotify complète souvent la recherche de l'utilisateur sans qu'il ait besoin de saisir le mot en entier.

### 2. Ajout de la visualisation avec les points d'écoute les plus intenses (Disjoint force-directed graph)

Une nouvelle visualisation a été ajoutée, reprise d'une visualisation exemple de D3 ([lien vers la visualisation](#)). Cette visualisation représente les articles sous forme de cercles rouges auxquels sont reliés des cercles jaunes représentant leurs morceaux écoutés par l'utilisateur. Cette visualisation permet de voir rapidement les artistes phares de l'utilisateur.

### 3. Dernières finitions

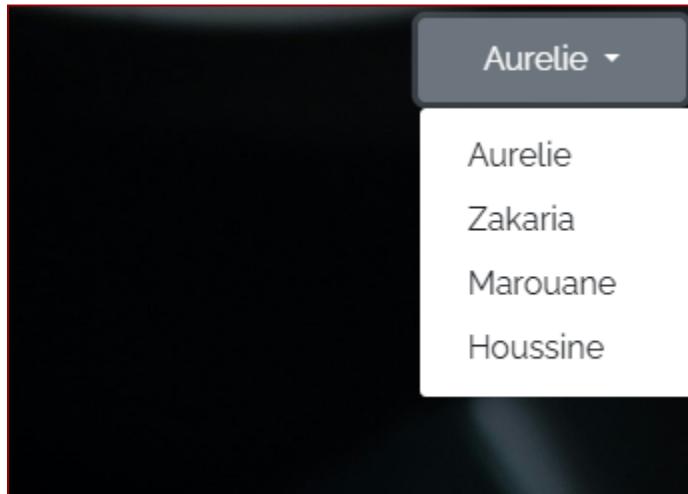
Les dernières finitions ont été ajoutées aux visualisations existantes : des titres ont été ajoutés aux axes de visualisations, les tooltips ont été ajustés et mis à jour, etc.

## Septième semaine : 14/01 - 21/01

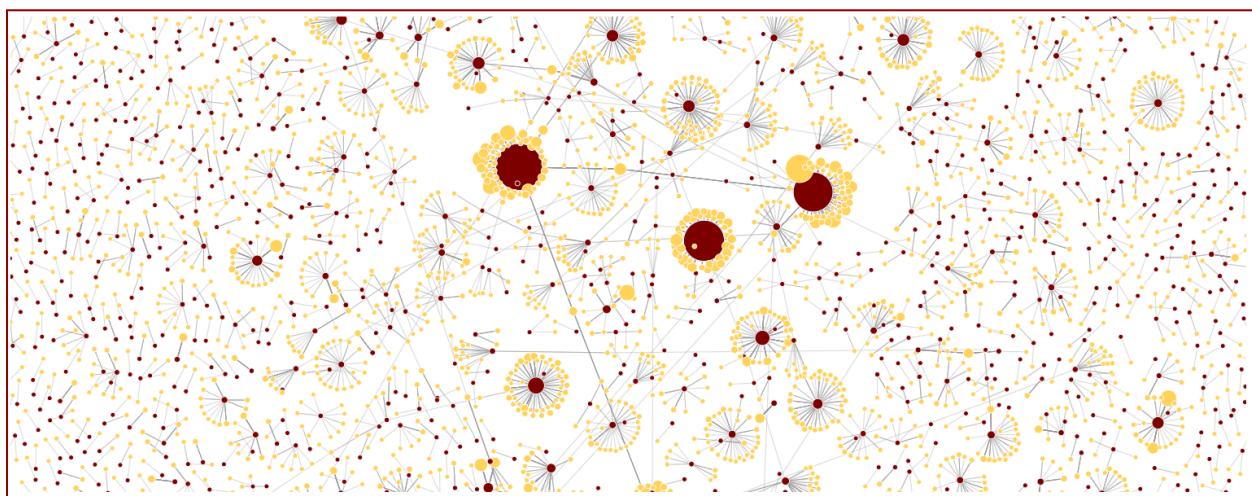
### 1. Finitions après la première présentation

Après la première présentation du projet, quelques finitions ont dû être apportées au projet grâce aux remarques constructives du professeur :

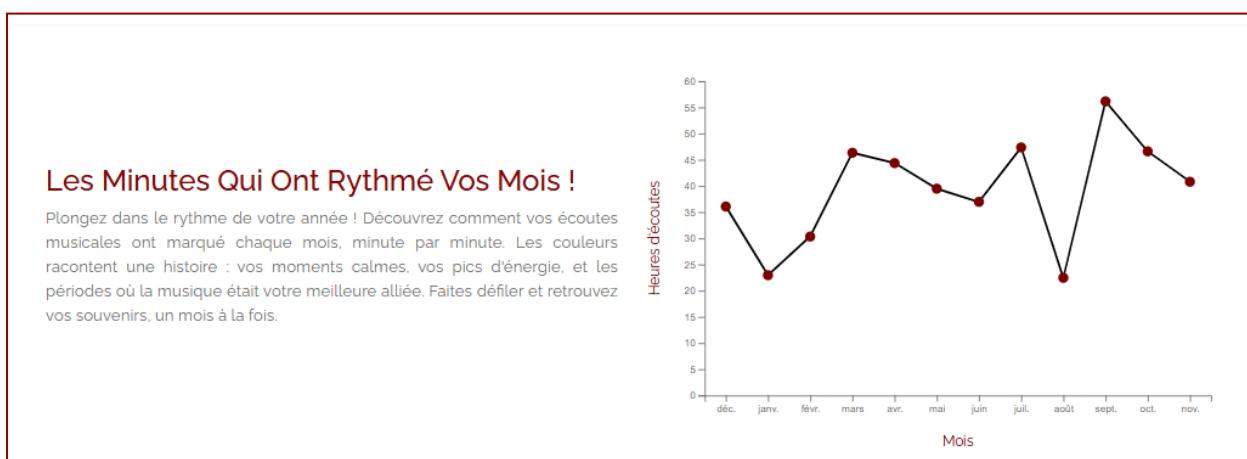
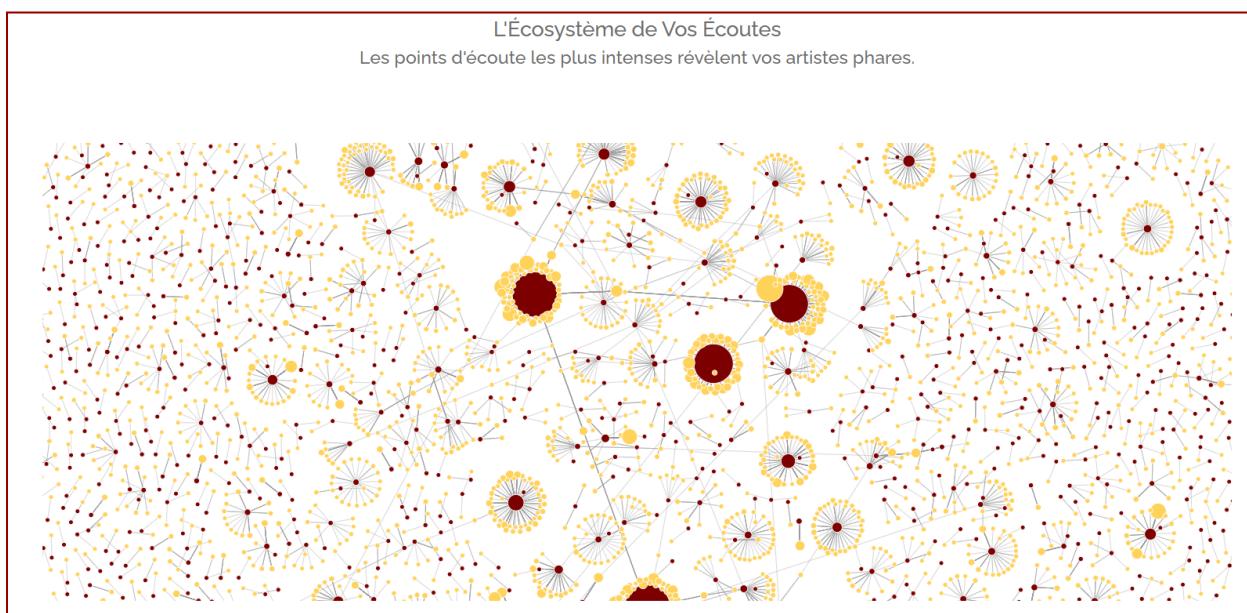
- Les noms d'utilisateur ont été mis à jour ("user-1", "user-2" sont devenus les prénoms des propriétaires des données)

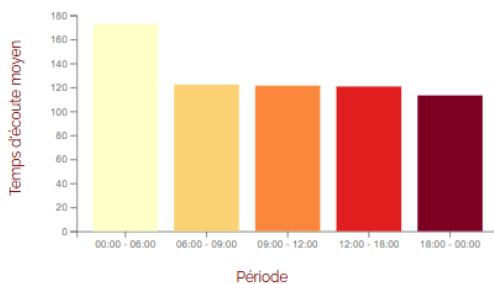


- Le traitement/chargement des données a été optimisée pour accélérer la performance du site
- Le filtre entre l'écoute de musique par mois et l'écoute moyenne par période de la journée était défectueux et a été fixé
- La visualisation inspirée de la visualisation exemple de D3 a été modifiée pour la personnaliser un peu plus par rapport à l'exemple existant : maintenant, la taille des cercles est proportionnelle au nombre d'écoutes de l'artiste ou du morceau



## Résultat final



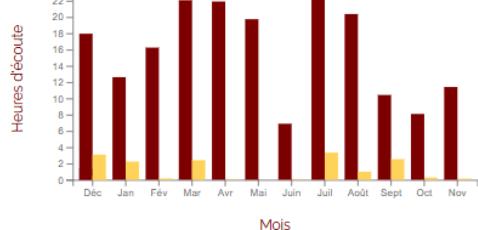
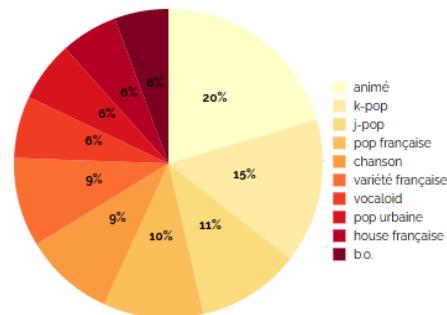


## Ton cycle quotidien d'écoute

La musique rythme tes journées : voici comment elle accompagne ton matin, ton après-midi, et tes soirées.

## Le genre musical qui te définit

Chaque note raconte une part de toi : découvre comment tes préférences musicales se répartissent et définissent ton univers sonore.



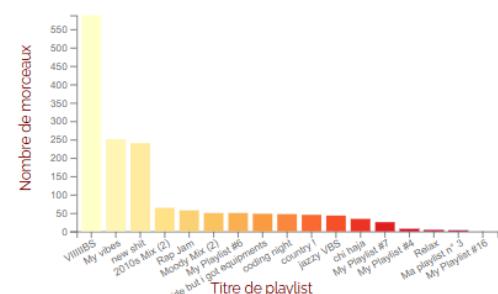
## Ton équilibre entre podcasts et musique

Plongé entre les rythmes de la musique et les récits captivants des podcasts, découvre comment tu partages ton temps entre ces deux univers sonores.

## Top 20 des Playlists les Plus Remplies : Un Océan de Musique

Découvrez le classement de vos 20 playlists les plus remplies de chansons ! Cette visualisation met en avant vos collections musicales les plus vastes, selon le nombre de morceaux qu'elles contiennent. Grâce à un filtre interactif, vous pouvez également explorer comment le nombre de chansons dans chaque playlist évolue mois après mois, révélant les tendances et les dynamiques musicales au fil de l'année.

Décembre 2024 ▾



### **Remarque :**

Les commits sur GitHub ne reflètent pas justement la contribution de chacun des membres du groupe (certains membres ont travaillé en local avant d'envoyer leur avancée à un autre membre du groupe qui l'a ensuite commit lui-même). Ce projet a été le fruit d'un travail collectif et bien réparti, et toutes les parties du projet ont bénéficié de la même contribution et des mêmes efforts de la part de tous les membres.